

Back-of-the-Envelope Miniload Throughput Bounds and Approximations

Robert D. Foley
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
rfoley@isye.gatech.edu

Steven T. Hackman
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332
shackman@isye.gatech.edu

Byung Chun Park
Department of Industrial Engineering
Keimyung University
Taegu, Korea
bcpark@kmu.ac.kr

June 14, 2002

Abstract

We derive simple formulas bounding and approximating the throughput of an end-of-aisle miniload system with exponentially distributed pick times and either uniform or turnover-based storage. For typical configurations, the worst case relative error for the bounds is less than 4%. We use our bounds to show that, for all practical purposes, *regardless* of the configuration, the picker utilization determines the s/r machine utilization, and vice-versa. Thus, a system designer cannot hope to independently achieve separate goals for the utilization rates.

Keywords: Miniload, automated storage/retrieval system, order picking, throughput, utilization, picker utilization, s/r machine utilization

1 Introduction

A miniload automated storage and retrieval system (AS/RS) can provide excellent space utilization, accurate picking and security for facilities that store and handle many thousands of small-parts. With such extremely high initial capital costs—a storage/retrieval (s/r) machine will cost several hundreds of thousands of dollars—throughput estimation is critical to the design of a miniload system.

Bozer and White (1990) developed approximate expressions for the throughput of end-of-aisle (EOA) order picking systems, assuming uniformly distributed activity and exponential or uniform pick time distributions. Bozer and White (1996) extended their work to EOA order picking systems with horse-shoe front-end configuration and/or two or more aisles per picker configuration (see also Park *et al.* 1999c).

Exact expressions for miniload throughput have been obtained for some special systems. Building on the early work of Hausman *et al.* (1976), Graves *et al.* (1977) and Bozer and White (1984), who analyzed unit-load AS/RS systems, Foley and Frazelle (1991) developed an exact expression for “square-in-time, uniformly distributed” rack. Their work has been extended to “rectangular-in-time” racks with uniform storage (Scharfstein 1991) and to square-in-time racks with 2-class storage (Park *et al.* 1999a). (See also Park *et al.* 1999b for “turnover-based” storage.) In almost all cases, the expressions are lengthy and cumbersome. Exact expressions are not available for many situations including turnover-based storage. Motivated by this, we derive simple analytical expressions suitable for “back-of-the-envelope” analysis to bound miniload throughput for an EOA miniload with uniform or turnover-based storage.

The outline of our paper is as follows. In Section 2, we briefly describe the class of miniload systems we analyze in this paper. In Section 3, we derive the bounds and approximations. In Section 4, we analyze a representative miniload system. In Section 5, we demonstrate the accuracy of our bounds. In Section 6, we show that a system planner cannot expect to independently set the picker and s/r machine utilization rates. Concluding remarks appear in Section 7.

2 Miniload AS/RS Background

A typical EOA miniload system consists of multiple aisles of storage rack, s/r machine operating within each aisle, modular storage containers for housing the items, and load stands at the end of each aisle to facilitate order picking. The load stands are arranged so that each aisle has 2 pick positions. While the order picker extracts items from the container in one pick position, the s/r machine picks up the container in the other pick position, travels to this container’s dedicated location within the rack, deposits the container, and then travels to the dedicated location of the next container to be retrieved, picks up the container, travels back to the load stand, and deposits this container in the open pick position. At this time, the s/r machine may have to wait to repeat the process if the order picker has not yet finished picking items from

the other container at the load stand. The s/r machine is said to execute a “dual command” cycle, as it handles 2 containers during each trip. (A single command cycle, commonly used for storing and handling unit loads, involves either a storage or a retrieval during a cycle, but not both.)

The s/r machine travels in the horizontal and vertical directions simultaneously. Consequently, the time to travel between 2 locations within the rack can be expressed as the maximum of the time required to traverse either the horizontal or vertical distance between these locations. We assume instantaneous acceleration and deceleration of the s/r machine travel, and that the sequence of retrieval locations are independent, identically distributed random locations in the rack.

The throughput of a miniload system is obviously highly dependent on the distribution of the dual command s/r machine cycle time, which may be significantly enhanced by stock assignment, namely, by assigning the most frequently requested containers to the closest locations in the rack. The success of this “turnover-based” storage policy is dependent on variation among the activity of containers, which we model with cumulative distribution function $F(x) = x^s$, $0 \leq x \leq 1$, $s > 0$. For example, if $s = 0.5$, then the top 20% of the most active containers will generate $.2^5 \approx 44.7\%$ of all picking activity. We use containers instead of stock keeping units (sku’s), since more than 1 sku may be housed in a container, and sku’s have different sizes and inventories. If the familiar 80-20 *Pareto’s law* holds, then we would have $s = \ln .8 / \ln .2 \approx 0.139$. The parameter “ s ” is referred to as the *skewness* of the distribution. In what follows when we speak of “uniformly distributed activity”, we mean that $s = 1$ and all containers are equally likely to be requested. (It might be better to call s , the “lack of skewness” since no skewness corresponds to $s = 1$, and skewness of the distribution increases as s decreases to 0; however, we will refer to s as the skewness since it agrees with the usage of several earlier papers in this area.)

Foley *et al.* (1999) showed that in many situations assuming exponentially distributed pick times provides a conservative bound on throughput. Specifically, they showed that among the class of NBUE (new better than used in expectation) with a given mean, the exponential distribution gives the lowest miniload throughput, while the deterministic distribution gives the highest throughput. In our context, a pick time distribution is NBUE if the expected remaining pick time is smaller than the expected pick time. Such distributions are frequently encountered in reliability and stochastic orderings; see Barlow and Proschan (1975) or Ross (1983). For the developments of our bounds to follow we assume pick times are exponentially distributed. Note that the throughput for another NBUE distribution, e.g., the Uniform, is bounded below by the exponential distribution we assume here, and bounded above by the deterministic case, for which closed-form throughput expressions exist for a variety of EOA systems.

3 Derivation of the Bounds and Approximations

Foley and Frazelle (1991) showed under more general assumptions than those of this paper that the throughput tp of a miniload system can be expressed as $1/m$ where $m = E[\max(P, C)]$. The random variables P and C are independent, P has the same distribution as a pick time, and $C \equiv c + D$ has the distribution as the dual command cycle time of the miniload system. The constant c accounts for the fixed time of moving containers off and on the s/r machine during a cycle and the variable portion D accounts for the round-trip travel time from the I/O point to the two random container locations and back. They have also shown that the picker utilization is given by $u_p = E[P]/m$ and the s/r machine utilization is given by $u_m = E[C]/m$.

Since the pick times are exponentially distributed with parameter λ , we have:

$$\begin{aligned}
 m &= E[C] + E[(P - C)^+] \\
 &= E[D] + c + E[(P - C)^+ | P \geq C] \Pr\{P > C\} \\
 &= E[D] + c + \Pr\{P > C\}/\lambda \\
 &= E[D] + c + E[e^{-\lambda C}]/\lambda \\
 &= E[D] + c + e^{-\lambda c} E[e^{-\lambda D}]/\lambda.
 \end{aligned} \tag{1}$$

Remark 1 We have assumed exponentially distributed pick times; however, the assumption can be relaxed. Let a denote the maximum dual command travel time. Note that

$$m = E[P | P > a] \Pr\{P > a\} + E[\max(P, C) | P \leq a] \Pr\{P \leq a\}.$$

Thus, as long as $E[P - a | P > a] = 1/\lambda$, then the pick time cdf only needs to be $1 - e^{-\lambda t}$ for $t \in [0, a]$, and all of our results still hold. The exact form of the cdf beyond a is irrelevant.

We now turn to bounding m . If we assume that we know the variable portion of the expected dual command travel time $E[D]$, then the only term in (1) that is not simple to evaluate is $E[e^{-\lambda D}]$. To develop our lower and upper bounds on throughput, we will develop upper and lower bounds on $E[e^{-\lambda D}]$, respectively. Since $e^{-\lambda D}$ is convex in D , application of the well-known Jensen's inequality yields the following lower bound on $E[e^{-\lambda D}]$:

$$e^{-\lambda E[D]} \leq E[e^{-\lambda D}]. \tag{2}$$

To find an upper bound, note that

$$\begin{aligned}
 E[e^{-\lambda D}] &= E[1 - \lambda D + \lambda^2 D^2/2! - \lambda^3 D^3/3! + \dots] \\
 &= 1 - \lambda E[D] + \lambda^2 E[D^2]/2! + E[-\lambda^3 D^3/3! + \dots] \\
 &= 1 - \lambda E[D] + \lambda^2 \{E[D]^2 + \text{Var}[D]\}/2! + E[-\lambda^3 D^3/3! + \dots] \\
 &\leq 1 - \lambda E[D] + \lambda^2 \{E[D]^2 + \text{Var}[D]\}/2! - \lambda^3 E[D]^3/3! + \dots \\
 &= e^{-\lambda E[D]} + \lambda^2 \text{Var}[D]/2
 \end{aligned} \tag{3}$$

where the inequality follows since the function $(-\lambda^3 D^3/3! + \dots)$ is concave.

With a little more work, we can obtain a better upper bound on $E[e^{-\lambda D}]$. The upper bound (3) is a convex function of λ with a global minimum at a point which we denote by λ_0 . Since $E[e^{-\lambda D}]$ is strictly decreasing in λ and is bounded above by (3), it then follows that

$$E[e^{-\lambda D}] \leq e^{-\tilde{\lambda}E[D]} + \tilde{\lambda}^2 \text{Var}[D]/2, \quad (4)$$

where $\tilde{\lambda} = \min\{\lambda, \lambda_0\}$.

The critical point λ_0 is the solution to

$$e^{-\lambda E[D]} = \frac{\text{Var}[D]}{E[D]} \lambda, \quad (5)$$

which may be expressed as

$$\lambda_0 = W\left(\frac{E[D]^2}{\text{Var}[D]}\right) / E[D]. \quad (6)$$

$W(x)$ is called Lambert's W -function, which gives the analytic solution y to the equation $ye^y = x$, and can be computed in mathematical packages such as *Maple* and *Mathematica*.

To summarize, we have $\underline{m} \leq m \leq \overline{m}$ and $\underline{tpt} \leq tpt \leq \overline{tpt}$ where

$$\underline{m} = E[D] + c + e^{-\lambda c} e^{-\lambda E[D]} / \lambda, \quad (7)$$

$$\overline{m} = E[D] + c + \frac{e^{-\lambda c}}{\lambda} \left\{ e^{-\tilde{\lambda}E[D]} + \frac{\tilde{\lambda}^2}{2} \text{Var}[D] \right\}, \quad (8)$$

$$\underline{tpt} = 1/\overline{m}, \quad (9)$$

$$\overline{tpt} = 1/\underline{m}. \quad (10)$$

Observe that the two throughput bounds converge to the exact throughput $1/(E[D] + c)$ as $\lambda \rightarrow \infty$.

We note that an easily derived two-moment approximation to miniload throughput is obtained by simply assuming that D is normally distributed. The moment generating function of a normally distributed random variable with mean $E[D]$ and variance $\text{Var}[D]$ is $e^{tE[D] + t^2 \text{Var}[D]/2}$. Using this gives:

$$\widehat{tpt} \equiv \left[E[D] + c + \frac{1}{\lambda} e^{-\lambda(c + E[D] + \lambda^2 \text{Var}[D]/2)} \right]^{-1} \approx tpt. \quad (11)$$

For large values of λ , \widehat{tpt} may fall below the lower bound \underline{tpt} . So we use $\widehat{tpt} \equiv \max\{\widehat{tpt}, \underline{tpt}\}$ as the *two-moment approximation*.

4 Analysis of a Representative System

We begin by describing reasonable ranges for the key variables describing a miniload system. We obtained these values from miniload system designers [5].

Table 1: Ranges for Miniload System Parameters

Name	Min	Max	Units
rack height	15	30	feet
rack length	30	300	feet
vertical speed	190	390	feet/minute
horizontal speed	300	750	feet/minute
total shuttle time/cycle	.267 (16)	.533 (32)	minutes (seconds)
average pick time/container	.167 (10)	1.0 (60)	minutes (seconds)
pick rate	60	360	containers/hour

We will give minimum and maximum values for many of the system parameters; however, simply because a particular set of parameters fall within the ranges given in the table, does not mean that someone has or would consider building such a system. In other words, some of the combinations might be unreasonable.

Table 1 summarizes the ranges for the height and length of the rack, the horizontal and vertical speeds of the s/r machine, the fixed shuttle cycle times, and average pick times (or pick rates). Each shuttle cycle (pick-up/deposit) takes 4-8 seconds so that 4 such cycles per dual command cycle will take 16-32 seconds. With respect to an average pick time a reasonable range is 5-30 seconds to pick a single sku. Usually, a picker will be picking more than 1 sku when processing a container: for example, for handling books one can expect roughly 1-2 sku's per tray; for handling parts, roughly 1-4 sku's per tray; and for assembly operations, roughly 1-10 sku's per tray. A reasonable range on the average pick rate is between 60 and 360 containers per hour.

We shall follow the common convention of converting to a time-normalized racks, which will also require determining the "shape factor" b of the rack (Bozer and White 1984). To do this, let T_V denote the time it takes the s/r machine to travel from the top to the bottom of the rack, let T_H denote the time it takes the s/r machine to travel the length of the aisle from the front to the back, and define the "time-normalizing factor" T as $\max(T_V, T_H)$. The rack shape factor b is defined as $\min(T_V, T_H)/T$. When $b = 1$, the rack is said to be square-in-time; otherwise, the rack is simply referred to as rectangular-in-time. Since the motors have significantly different speeds, a square-in-time rack will not be physically square.

To compute throughput, we convert time data from minutes to the normalizing time units. From the values in Table 1, the time-normalizing factor T will lie between .04 and 1.0 minutes. The total shuttle time per cycle c will lie between .267 and 13.3 normalized time units, and the pick rate λ will lie between .04 to 6.0 containers per normalized time unit. Note that some of these values are unreasonable since they arise from unreasonable combinations of values from

Table 2: Some Characteristic Values for Time-Normalized Racks with Turnover-Based Storage

		skewness parameter " s "							
		1.0 (20%)	0.748 (30%)	0.569 (40%)	0.431 (50%)	0.317 (60%)	0.222 (70%)	0.139 (80%)	0.065 (90%)
$E[D]$	1.0	1.8000	1.6628	1.5211	1.3679	1.1923	0.9892	0.7403	0.4204
	0.8	1.6363	1.5093	1.3789	1.2384	1.0781	0.8934	0.6678	0.3788
	0.6	1.5061	1.3834	1.2584	1.1254	0.9755	0.8049	0.5990	0.3382
	0.4	1.4112	1.2873	1.1626	1.0319	0.8871	0.7257	0.5351	0.2991
	0.2	1.3531	1.2239	1.0952	0.9619	0.8170	0.6590	0.4780	0.2621
	0.0	1.3333	1.1987	1.0646	0.9259	0.7760	0.6150	0.4351	0.2301
$\text{Var}[D]$	1.0	0.1767	0.2193	0.2651	0.3125	0.3581	0.3912	0.3906	0.3022
	0.8	0.1581	0.1944	0.2319	0.2698	0.3053	0.3297	0.3258	0.2497
	0.6	0.1661	0.2011	0.2336	0.2631	0.2877	0.3003	0.2870	0.2132
	0.4	0.1882	0.2273	0.2595	0.2842	0.2994	0.2994	0.2726	0.1920
	0.2	0.2114	0.2580	0.2948	0.3200	0.3307	0.3207	0.2793	0.1854
	0.0	0.2222	0.2747	0.3173	0.3475	0.3606	0.3485	0.2989	0.1912
λ_0	1.0	1.1920	1.1397	1.0912	1.0463	1.0048	0.9693	0.9430	0.9379
	0.8	1.2782	1.2240	1.1756	1.1311	1.0902	1.0554	1.0302	1.0278
	0.6	1.2931	1.2391	1.1960	1.1597	1.1281	1.1031	1.0878	1.0953
	0.4	1.2625	1.2033	1.1613	1.1306	1.1084	1.0950	1.0935	1.1158
	0.2	1.2231	1.1546	1.1062	1.0720	1.0487	1.0373	1.0407	1.0684
	0.0	1.2044	1.1284	1.0719	1.0283	0.9946	0.9712	0.9590	0.9640

Table 1. In particular, typical values for c would be one or larger.

To use the expressions from Section 3, we need the mean $E[D]$ and variance $\text{Var}[D]$ of the dual command travel time as well as the critical point λ_0 . Table 2 reports values for $E[D]$, $\text{Var}[D]$ and λ_0 for time-normalized racks with turnover-based storage. The results for $E[D]$ and $\text{Var}[D]$ are reproduced from Park *et al.* (1999b).

We analyze the following miniload, which is a reasonable configuration according to the system designers [5]. The rack is 24 feet high and 150 feet long. The s/r machine travels 240 FPM in the vertical direction and 600 FPM in the horizontal direction. The I/O point is located at the lower left-hand corner of the rack. The (total) constant pickup/deposit time is .4 minutes (25 seconds). Finally, the picker's average pick time per container is 3/4 minutes (45 seconds), i.e., the pick rate is 80 containers per hour.

Since $\max\{150/600, 24/240\} = 0.25$, the time-normalizing factor $T = 0.25$ minutes while the rack shape factor $b = (24/240)/(150/600) = 0.4$. In normalized time, the pick rate $\lambda = T/(3/4) = 1/3$, and $c = .4/T = 1.6$.

Consider first the case when the skewness parameter $s = 1.0$, i.e., uniform stock assignment. From Table 2 we see that $E[D] = 1.4112$, $\text{Var}[D] = 0.1882$ and $\lambda_0 = 1.2625$. Since $\lambda_0 > \lambda$, $\tilde{\lambda} = \lambda = 1/3$. Substitute these values in the expressions for the bounds on m in (7) and (8), and then use these values in (9) and (10) to obtain the lower bound 0.2422 and the upper bound 0.2433 on the throughput in normalized time units. By dividing these values by T , the throughput lower bound is 0.9689 containers per minute or 58.13/hr, and the upper bound is 0.9732 containers per minute or 58.39/hr. Note that the upper and lower bounds imply that the relative error must be less than 1/2 of one

percent; the bounds are exceptionally tight.

The two-moment approximation yields 0.2426 containers per normalized time unit or 58.22 containers per hour. Since the average pick rate is 80 containers per hour, the picker utilization is $58.22/80 = 73\%$. Since the average dual command cycle time is $1.4112 + 1.6 = 3.0112$ time units or 0.7528 minutes, which equates to 79.70 containers per hour, the s/r machine utilization is $58.22/79.70 \approx 73\%$.

Now suppose there is a high degree of stock assignment planning so that the most active 20% of the containers generate 80% of the total container activity. Here, the skewness parameter $s = 0.139$. From Table 2, we see that $E[D] = 0.5351$ and $\text{Var}[D] = 0.2726$. (Once again λ_0 plays no role.) Thus, the lower bound = 1.1008 containers per minute or 66.05/hr; the two-moment approximation = 1.1019 containers per minute or 66.12/hr; and the upper bound = 1.1088 containers per minute or 66.53/hr. The picker utilization in this case is 83% and the s/r machine utilization is 59%. The bounds imply that the relative error is less than one percent; the bounds are once again exceptionally tight.

5 Performance of Bounds and Approximations

To see how accurate the throughput bounds and approximations are, we look at the worst case relative error where the relative error is given by

$$\frac{\overline{tpt} - \underline{tpt}}{\underline{tpt}} = \frac{\overline{m}}{\underline{m}} - 1. \quad (12)$$

The worst case relative error is particularly sensitive to the value of c . It is not hard to show analytically that the worst case relative error increases as c decreases. The minimum possible value for c obtained using the values in Table 1 is $c = .267$. Table 3 contains the worst case relative error for each of the cases given in Table 2 assuming $c \geq .267$. The number in parentheses below the relative error bound is the value of λ that gives the worst value for the relative error bound for the specified configuration. Note that in all cases the relative error is smaller than 14%. If we estimate the throughput by using the midpoint of the throughput bounds, the relative error must be smaller than 7% (though in practice the two-moment approximation seems to perform better than the midpoint).

Typical configurations have $c \geq 1$. Table 4 contains the same information as the Table 3 except $c \geq 1$. For these systems, the worst case relative error is less than 4%. Again the midpoint approximation would have a worst case relative error of half the size, i.e., 2%.

Note that relative error bounds appearing in Table 3 and Table 4 are substantially higher than actual relative errors usually encountered. For example, in the system analyzed in the previous section with $b = 0.4$, $c = 1.6$ and $\lambda = 1/3$, the relative error of the bound is 0.45% for $s = 1$ and 0.73% for $s = 0.139$, while the corresponding numbers in Table 4 are 1.39% and 2.94%, respectively.

Table 3: Maximum Relative Error Bounds ($c \geq .267$)

		skewness parameter " s "							
		1.0 (20%)	0.748 (30%)	0.569 (40%)	0.431 (50%)	0.317 (60%)	0.222 (70%)	0.139 (80%)	0.065 (90%)
b	1.0	3.8% (1.38)	4.9% (1.35)	6.1% (1.33)	7.6% (1.33)	9.3% (1.34)	11.3% (1.38)	13.2% (1.49)	13.7% (1.80)
	0.8	3.8% (1.48)	4.9% (1.45)	6.1% (1.43)	7.5% (1.43)	9.1% (1.44)	10.9% (1.49)	12.6% (1.60)	12.8% (1.91)
	0.6	4.4% (1.51)	5.5% (1.48)	6.7% (1.47)	8.0% (1.48)	9.4% (1.51)	11.0% (1.57)	12.3% (1.70)	12.0% (2.01)
	0.4	5.1% (1.50)	6.4% (1.47)	7.7% (1.47)	9.0% (1.49)	10.3% (1.53)	11.6% (1.61)	12.5% (1.76)	11.6% (2.10)
	0.2	5.8% (1.48)	7.4% (1.45)	8.8% (1.45)	10.2% (1.47)	11.5% (1.52)	12.6% (1.61)	13.0% (1.79)	11.3% (2.14)
	0.0	6.1% (1.46)	7.8% (1.43)	9.4% (1.43)	11.0% (1.45)	12.4% (1.51)	13.4% (1.61)	13.6% (1.79)	11.2% (2.16)

Table 4: Maximum Relative Error Bounds ($c \geq 1.0$)

		skewness parameter " s "							
		1.0 (20%)	0.748 (30%)	0.569 (40%)	0.431 (50%)	0.317 (60%)	0.222 (70%)	0.139 (80%)	0.065 (90%)
b	1.0	1.14% (1.07)	1.48% (1.10)	1.88% (1.09)	2.35% (1.06)	2.87% (1.03)	3.38% (1.02)	3.73% (1.02)	3.34% (1.05)
	0.8	1.08% (1.08)	1.39% (1.09)	1.74% (1.10)	2.13% (1.11)	2.57% (1.10)	3.00% (1.08)	3.27% (1.08)	2.88% (1.11)
	0.6	1.19% (1.09)	1.50% (1.10)	1.83% (1.11)	2.18% (1.13)	2.54% (1.13)	2.85% (1.12)	3.00% (1.13)	2.54% (1.16)
	0.4	1.39% (1.10)	1.76% (1.11)	2.12% (1.12)	2.45% (1.13)	2.74% (1.12)	2.95% (1.12)	2.94% (1.14)	2.34% (1.18)
	0.2	1.60% (1.10)	2.05% (1.11)	2.47% (1.11)	2.83% (1.09)	3.11% (1.08)	3.24% (1.09)	3.07% (1.11)	2.29% (1.16)
	0.0	1.69% (1.10)	2.21% (1.12)	2.69% (1.09)	3.11% (1.06)	3.43% (1.05)	3.55% (1.05)	3.30% (1.06)	2.34% (1.10)

Typically, the relative error for the two-moment approximation is also substantially smaller than the bound given in Table 3 and Table 4. For example, in square-in-time, uniform case, the bound on the relative error given in Table 4 is 1.14%. However, the relative error $(\widehat{tpt} - tpt)/tpt$ for the two-moment approximation in the same case is always less than .7%, for all values of $c \geq 0$ and $\lambda > 0$.

Note that the maximum error bound is not monotonic in b or s . The maximum error bound does seem to behave similarly as the variance in Table 2.

6 A Utilization Rate "Identity"

In this section, we show that the picker utilization essentially determines the s/r machine utilization, and vice-versa. For example, if the picker utilization is 95%, then the s/r machine utilization must be roughly 33%. On the other hand, if the s/r machine utilization is 95%, the picker utilization will be roughly 44%.

Thus, it is *impossible* for the system planner to achieve separate goals for the picker and s/r machine utilization rates.

The utilization of the picker is given by $u_p = 1/(\lambda m)$, and the utilization of the s/r machine is $u_m = (E[D] + c)/m$. Thus, we have $u_m = \theta u_p$ where $\theta = \lambda(E[D] + c)$ denotes the ratio of the expected dual command s/r machine cycle time to the expected pick time.

Since $\underline{m} \approx m$, we obtain the following simple approximations to the utilizations:

$$\hat{u}_p \equiv \frac{1}{\lambda(E[D] + c) + e^{-\lambda(E[D] + c)}} = \frac{1}{\theta + e^{-\theta}} \approx u_p \quad (13)$$

$$\hat{u}_m \equiv \frac{\lambda(E[D] + c)}{\lambda(E[D] + c) + e^{-\lambda(E[D] + c)}} = \frac{\theta}{\theta + e^{-\theta}} \approx u_m, \quad (14)$$

which are shown in Figure 1. Observe that both relative errors $\hat{u}_p/u_p - 1$ and $\hat{u}_m/u_m - 1$ equal the relative error $m/\underline{m} - 1$, which we have demonstrated to be quite small. Thus, θ essentially determines u_p and u_m . In fact, \hat{u}_p and \hat{u}_m are also upper bounds since $\underline{m} \leq m$, and they can be combined to yield the following “identity” (which we put in quotes since they involve \hat{u}_p and \hat{u}_m instead of the actual utilizations):

$$\hat{u}_m + \hat{u}_p e^{-\hat{u}_m/\hat{u}_p} = 1$$

and is shown in Figure 2.

Note that we can directly obtain the throughput estimate \overline{tpt} by multiplying the picker utilization \hat{u}_p by the pick rate. In our representative example with uniformly distributed activity, i.e., $s = 1$, the pick rate was 80 containers per hour and $\hat{u}_p \approx 73\%$ implying that $\overline{tpt} \approx 58$ containers per hour.

At the beginning of this section, we considered an example in which the picker utilization was 95%. This implies that $\theta \approx .34$, which means that the s/r machine utilization is roughly 33%. The most we can possibly obtain if we attempt to maximize the minimum of \hat{u}_p and \hat{u}_m is 73% utilization for both.

7 Concluding Remarks

Since the upper and lower throughput bounds are so close, it is reasonable to use the upper bound \overline{tpt} to approximate tpt . The upper bound has the advantage that it only depends on the distribution of D through its mean $E[D]$, which can be found using Table 2 or estimated.

The expression for the upper throughput bound makes sensitivity analysis easy. For example, suppose that a user of a miniload system is considering converting from uniform storage to class-based storage. The upper throughput bound suggests choosing the partitions to minimize $E[D]$. It is easy to estimate the percentage improvement from the expression for the upper throughput bound. In particular, if λ is small relative to $E[D]$, the increase in throughput may not be sufficient to justify changing storage policies. On the other hand,

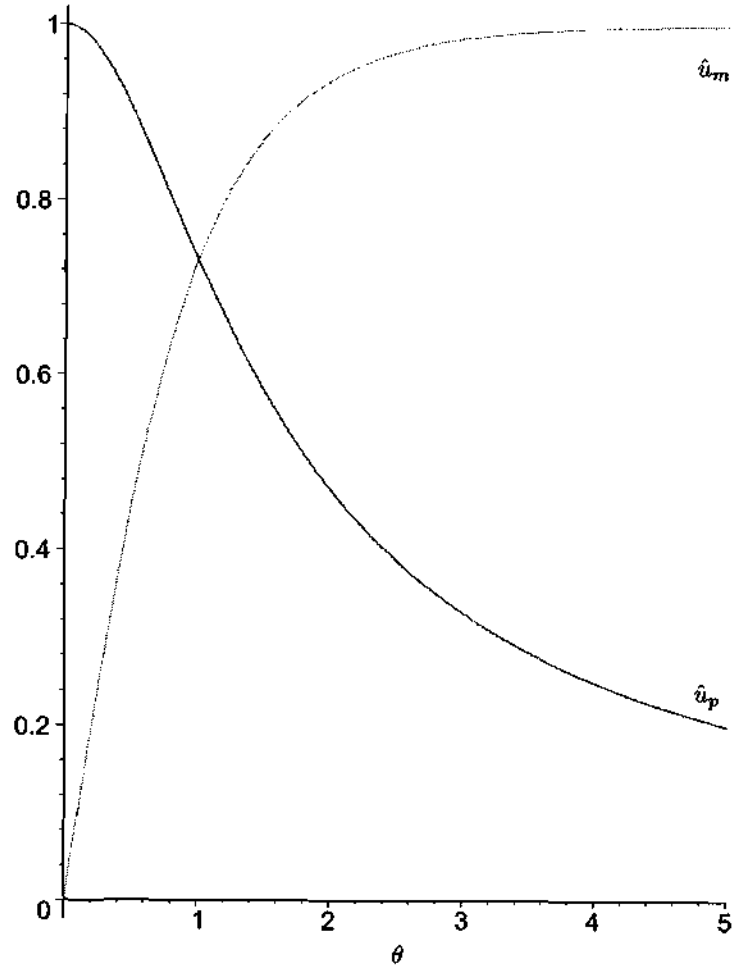


Figure 1: The utilizations \hat{u}_p and \hat{u}_m as a function of θ

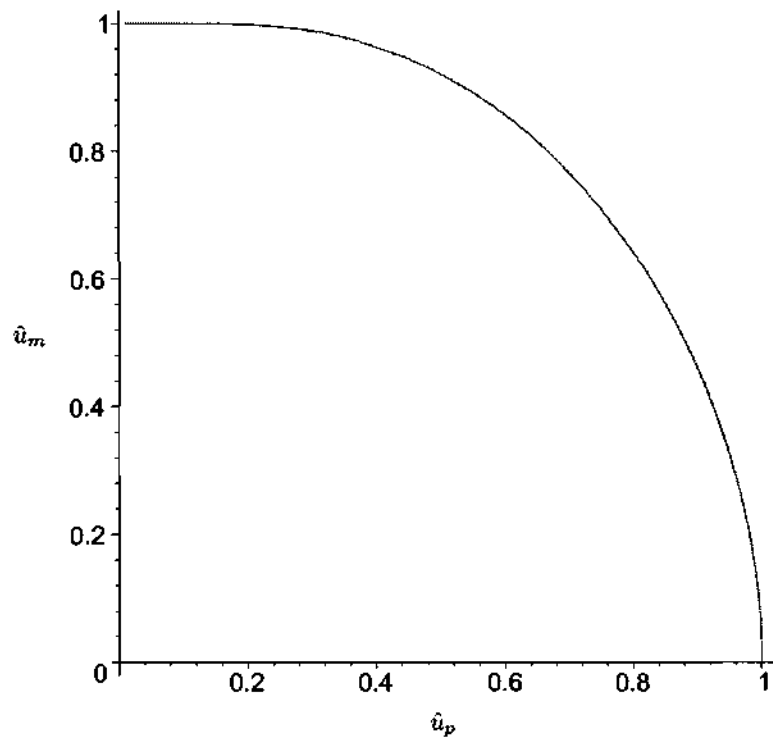


Figure 2: The utilization “identity” of \hat{u}_p and \hat{u}_m

if λ is large, then even a small decrease in $E[D]$ can yield a large increase in throughput.

Estimation of throughput and utilization rates of a miniload system with uniform or turnover-based storage can be reduced to the following simple steps:

1. Determine the time-normalizing factor T (in minutes), rack shape parameter b , and skewness s . Using T , compute the shuttle time parameter c and the pick rate λ .
2. Use Table 2 to obtain $E[D]$, $\text{Var}[D]$ and λ_0 .
3. Either use the results in Section 3 (more complete) to obtain bounds and approximations to throughput and utilizations, or compute θ and use the results in Section 6 (simpler) to estimate throughput and utilizations.

We have not discussed class-based storage systems, though their performance would lie between the performance of the uniform and turnover-based storage systems. If class-based storage were used instead of turnover-based storage in the representative example analyzed in Section 4, the results in this paper would imply that the throughput would be between 58.13 and 66.53 containers per hour. If this range were too wide for design purposes, the system planner should look at results specifically developed for class based storage. Of course, the results in this paper would apply to class-based storage systems if $E[D]$ and $\text{Var}[D]$ were known. Expressions for $E[D]$ and $\text{Var}[D]$ for square-in-time racks with 2-class storage can be found in Park *et al.* (1999a). For more than 2 classes or for rectangular class-based storage systems, we suggest either to bound $E[D]$ by considering better and worse systems with known results or to estimate $E[D]$ through simulation and then use the upper bound as an approximation as described in this paper.

Acknowledgements

The authors thank Richard Flynn, Roger Stubbs and Don Vanoostveen of Siemens-Dematic for their assistance in providing information for Section 4.

References

- [1] Barlow, R. E. and Proschan, F. (1975) *Statistical Theory of Reliability and Life Testing*, Holt, New York.
- [2] Bozer, Y. A. and White, J. A. (1984) Travel-Time Models for Automated Storage/Retrieval Systems. *IIE Transactions*, 16(4), 329-338.
- [3] Bozer, Y. A. and White, J. A. (1990) Design and Performance Models for End-of-Aisle Order Picking Systems. *Management Science*, 36(7), 852-866.

- [4] Bozer, Y. A. and White, J. A. (1996) A Generalized Design and Performance Analysis Model for End-Of-Aisle Order Picking Systems. *IIE Transactions*, 28(4), 271-280.
- [5] Flynn, R., Stubbs, R. and Vanoostveen, D. of Siemens-Dematic (2001). Personal communication.
- [6] Foley, R. D. and Frazelle, E. H. (1991) Analytical Results for Miniload Throughput and the Distribution of Dual Command Travel Time. *IIE Transactions*, 23(3), 273-281.
- [7] Foley, R. D., Frazelle, E. H. and Park, B. C. (1999) Throughput Bounds for Miniload Automated Storage/Retrieval Systems. Submitted for Publication.
- [8] Graves, S. C., Hausman, W. H. and Schwarz, L. B. (1977) Storage-Retrieval Interleaving in Automatic Warehousing Systems. *Management Science*, 23(9), 935-945.
- [9] Hausman, W. H., Schwarz, L. B. and Graves, S. C. (1976) Optimal Storage Assignment in Automatic Warehousing Systems. *Management Science*, 22(6), 629-638.
- [10] Park, B. C., Foley, R. D. and Frazelle, E. H. (1999a) Dual Command Travel Time Distribution and Performance of Miniload Systems with 2-Class Storage. Submitted for Publication.
- [11] Park, B. C., Foley, R. D., White, J. A. and Frazelle, E. H. (1999b) Dual Command Travel Times and Miniload System Throughput with Turnover Based Storage. Submitted for Publication.
- [12] Park, B. C., Frazelle, E. H. and White, J. A. (1999c) Buffer Sizing Models for End-of-Aisle Order Picking Systems. *IIE Transactions*, 31(1), 31-38.
- [13] Ross, Sheldon M. (1983) *Stochastic Processes*, John Wiley & Sons, New York, N.Y.
- [14] Scharfstein, D. O. (1991) Analytical Performance Measures for the Miniload Automated Storage/Retrieval System. Unpublished M.S. Thesis, Georgia Institute of Technology, Atlanta, Georgia.

Slotting Multi-Tier Inventory Systems

Stephanie Jernigan
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205
jernigan@isye.gatech.edu

John J. Bartholdi, III
Faculty Advisor
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205
john.bartholdi@isye.gatech.edu

November 16, 2003

1 Introduction

The most basic type of inventory system a warehouse can have is a single storage mode: each stock-keeping unit, or *sku* (pronounced “skew”), is picked from that mode, and replenishment stock is stored there after it is received into the warehouse. An example is a warehouse that has only pallet rack—even if a sku has a only single case present in the warehouse, the case is placed on a pallet and stored in a pallet location.

A warehouse that picks some skus with an automatic picking device may have a more complicated inventory system, as shown in Figure 1. The automatic device holds a small quantity of each sku, and this supply is replenished from a supply stored in case flow rack located nearby. We say that the automatic picker is *restocked* from the case flow rack. The supply in the case flow rack is replenished from bulk storage: this path is shown by the solid arrows in Figure 1. Skus in the warehouse that are not picked from the automatic picking device are picked from shelving that is restocked from bulk storage, as shown by the dotted arrows in Figure 1. This is an example of a multi-tier inventory system. Observe that no skus are picked

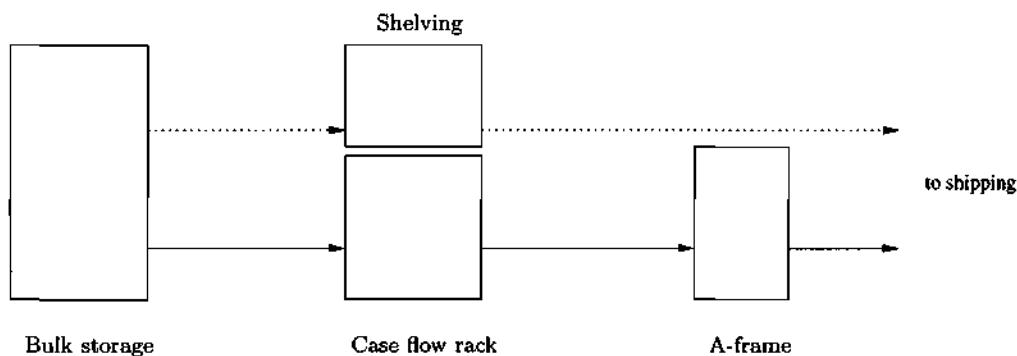


Figure 1: A multi-tier inventory system.

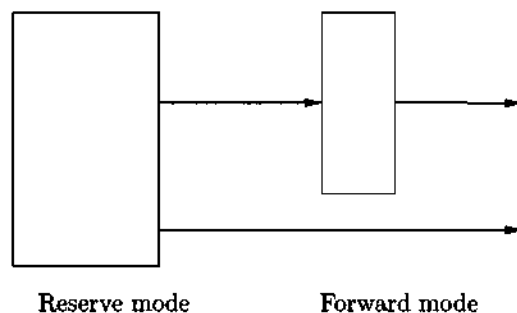


Figure 2: A forward-reserve inventory system

from the case flow rack or bulk storage; these storage modes serve only to restock the A-frame.

Definition 1 A multi-tier inventory system is a set of storage modes in a warehouse where associated with each storage mode is

- a set of modes that replenish it
- a designation of whether or not skus can be picked from the mode.

In the inventory system in Figure 1, for example, the storage mode consisting of case flow rack is replenished from bulk storage only, and skus can be picked from the mode. The commonly known forward-reserve inventory system, shown in Figure 2, is another example of a multi-tier inventory system. In this case, there are two storage modes: a reserve mode where skus are typically stored in bulk quantities, and a forward mode where skus are typically stored in smaller quantities. Skus may be picked from either the

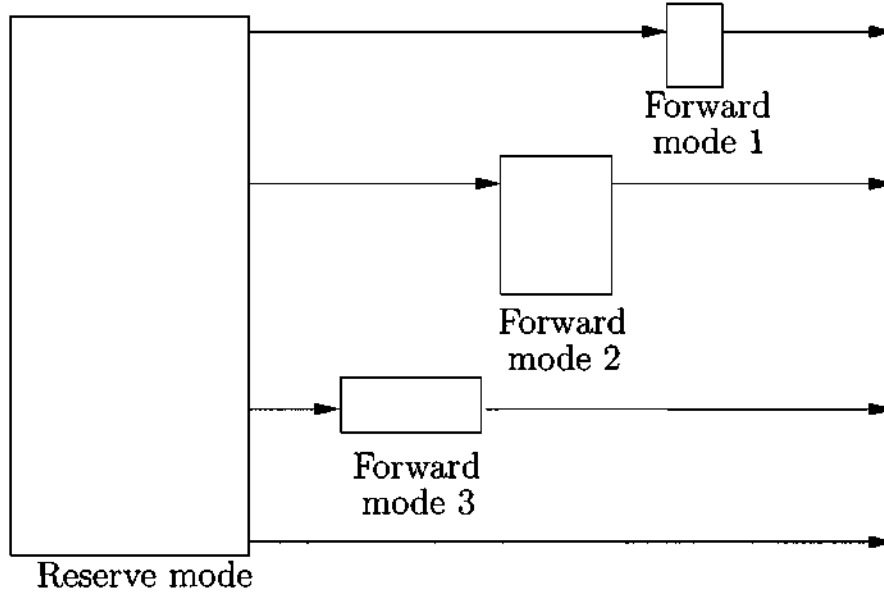


Figure 3: A multi-mode inventory system

forward mode or the reserve mode, and the supply of skus in the forward mode is replenished from the reserve mode. The multi-mode inventory system, an extension of the forward-reserve system with several forward modes, is also a multi-tier inventory system. As shown in Figure 3, skus may be picked from either the reserve mode or one of several forward modes, and the supply of skus in each forward mode is replenished from the reserve mode.

Storage modes in both the forward-reserve and multi-mode systems are restocked directly from a bulk storage mode, but this is not the case in other multi-tier inventory systems, such as the inventory system in the warehouse of Avon Products, Inc. near Atlanta.

2 Focus of this research

In a typical warehouse, 70% of the total operating cost is attributed to picking and restocking activities ([2]). To minimize these costs in an existing multi-tier inventory system, the warehouse manager can control which skus are stored in which storage modes, and how much of each sku is stored there. The process of deciding the storage modes to which to assign a sku and how much space to allot to the sku there is called *slotting* the inventory system.

Any result of this process is also referred to as a slotting. In some research, slotting the inventory system refers to assigning a bin location to each sku, but we use the term to mean only the storage modes to which a sku is assigned. We focus on slotting a multi-tier inventory system to minimize the cost of picking and restocking.

3 Model

3.1 Storage modes

Our model of a multi-tier inventory system is based on the model of the forward-reserve system developed by Hackman and Rosenblatt in [4]. We assume that every multi-tier inventory system that we model will have the following characteristics:

1. Upon receipt into the warehouse, all skus are stored in a bulk storage mode which has sufficient supply of each sku to restock all other modes as needed. This mode, called the reserve mode, is denoted Mode R .
2. Each storage mode in the system has a known storage capacity, denoted V_m for mode m . The reserve mode can be assumed to have infinite capacity. Our default unit of measurement will be cubic feet.
3. Associated with each storage mode is a set of storage modes that can supply restocks for the mode, called the *predecessor modes* for the mode.
 - In the example in Figure 4, the only predecessor mode for Mode 2 is Mode R . Mode 3 counts both Mode 2 and Mode R as predecessor modes.

The cost per restock of a sku in mode m from predecessor mode ℓ is denoted $c_{\ell m}$. The cost is the same for all skus restocked in mode m from mode ℓ . The restock cost is assumed to be nonnegative and is independent of the size of the restock. We do not consider the cost of restocking the reserve mode in this research.

If skus can be picked from mode m , the cost per pick is denoted d_m . We initially assume that this cost is the same for all skus picked from mode m , but we will note when we relax this assumption.

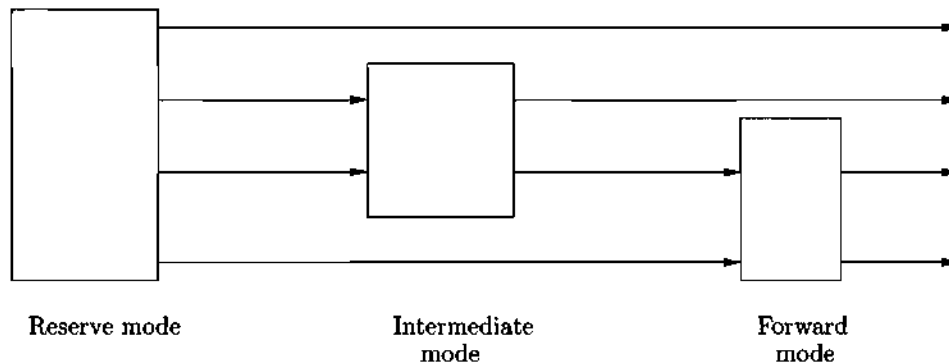


Figure 4: A three-tier inventory system.

3.2 Skus

Let S represent the set of skus to be stored in the inventory system, and assume that we are analyzing warehouse operations over a fixed period of time. For each sku $i \in S$, we know

- the total number of orders on which the sku appears each period, called the *picks* of sku i and denoted p_i , and
- the total cubic feet sold per period, called the *flow* of sku i and denoted f_i .

We will frequently use the square root of the flow of each sku in computations; we refer to the quantity $\sqrt{f_i}$ as the *rootflow* of sku i .

3.3 How to characterize a slotting

In the Hackman-Rosenblatt model of a forward-reserve inventory system, a slotting is characterized by stating which skus are picked from each mode and how much space each sku is allotted in each mode. In a multi-tier inventory system, however, this is not sufficient, since two skus picked from the same mode may have been restocked in that mode from different predecessor modes. In a slotting of a multi-tier inventory system, each sku is restocked in the mode from which it is picked via a path of storage modes, which we will call the *flowpath* of the sku. Each mode in the flowpath restocks its successor in the path.

When it is important to know the component nodes of a flowpath, we will use path notation from graph theory and denote each flowpath as a path

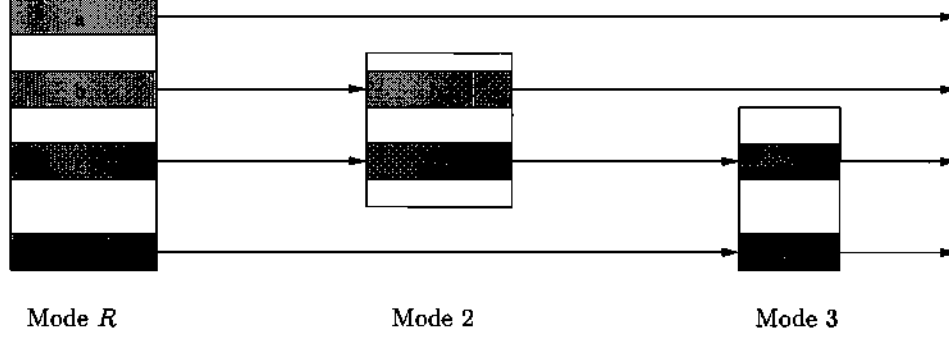


Figure 5: Skus assigned to the four possible flowpaths in a three-tier system.

of its component modes. Figure 5 shows a multi-tier inventory system with four flowpaths; skus a through d are each assigned to a different flowpath.

- Sku a is assigned to flowpath R and is picked from the reserve mode.
- Sku b is assigned to flowpath $R, 2$: it is picked from Mode 2, and its supply there is replenished from the reserve mode.
- Sku c is assigned to flowpath $R, 2, 3$: it is picked from Mode 3; its supply in Mode 3 is replenished from Mode 2; and its supply in Mode 2 is replenished from the reserve mode.
- Sku d is assigned to flowpath $R, 3$: it is picked from Mode 3; its supply in Mode 3 is replenished directly from the reserve mode.

We can characterize a slotting of a multi-tier inventory system by stating the assignment of skus to flowpaths and the amount of space that each sku is allotted in each mode in its flowpath. We make the following assumptions about the flowpaths in an inventory system.

1. Each sku is assigned to exactly one flowpath.
2. All flowpaths begin with the reserve mode.

The q th mode in the flowpath is referred to as the q th *tier* of the flowpath. A given storage mode may be the q th tier of one flowpath and the r th tier of another. In Figure 5, Mode 3 is the third tier of flowpath $R, 2, 3$ but the second tier of flowpath $R, 3$.

A flowpath can be thought of as a sequence of tiers, where tier q restocks tier $q + 1$. The *length* of a flowpath is defined as the number of modes it

comprises, and an inventory system whose longest flowpath has length ℓ is referred to as a ℓ -tier inventory system. The set of all skus assigned to a flowpath is known as a *flowgroup*; the flowgroup that corresponds to flowpath g is denoted $S(p)$.

3.4 Restocking protocol

Over the long term, the rate at which each sku is restocked in its picking mode must equal the rate at which the picking mode requests restocks of that sku. For analytical tractability, we make two simplifying assumptions. We assume that restocks of a sku in a mode occur instantaneously from the appropriate predecessor tier and that a sku is restocked in a storage mode when the supply of the sku there has been completely depleted, with no allowance for safety stock. We also assume that each storage mode can accumulate enough supply to completely restock a successor mode when needed, possibly through multiple restock events.

If v_{im} is the amount of space allotted to sku i in mode m , then sku i will be restocked in mode m a total of f_i/v_{im} times.

4 Previous research

This research to minimize the cost of picking and restocking multi-tier inventory systems builds on the results from previous research.

The model of a multi-tier inventory system is based on the forward-reserve inventory system presented by Hackman and Rosenblatt in [4], which is pictured in Figure 2. Hackman and Rosenblatt first assume an assignment of skus to the forward mode, and they show how to optimally allocate the available space among the skus. If $S(F)$ represents the set of skus assigned to the forward mode, then they show that

Theorem 1 *In order to minimize the cost of restocking the forward mode, sku i in flowgroup $S(R, F)$ should be allotted*

$$\frac{\sqrt{f_i}}{\sum_{j \in S(F)} \sqrt{f_j}} V_F \quad (1)$$

cubic feet in the forward mode.

Bartholdi and Hackman extended the Hackman-Rosenblatt model of a forward-reserve system to the case when there is more than one forward mode, which is referred to as the multi-mode inventory system, pictured in

Figure 3. The cost each time a sku is picked from a given storage mode is the same for all skus stored there, as is the cost each time a sku is restocked in a given forward storage mode. In addition, as in the Hackman-Rosenblatt forward-reserve system, skus are allotted a fraction of the total space in the forward modes, and it is assumed that they can fully occupy the space.

Bartholdi and Hackman rank the storage modes in descending order of cost per pick and rank the skus according to the same measure as in [4]. They show that a near-optimal assignment of skus to forward modes has the property that some number of the highest-ranked skus are assigned to the highest-ranked storage mode, and some number of the next highest-ranked skus are assigned to the next highest-ranked storage mode, and so on. This observation leads to the following theorem:

Theorem 2 *A provably near-minimum cost slotting of a multi-mode system can be found in $O(n)^M$ time by considering only those solutions that correspond to a partition of the skus ranked by viscosity.*

Bartholdi and Hackman also derive properties of the objective function that allow them to find a near-minimum cost slotting of a multi-mode system in $O(\log n^M)$ time.

The research on the multi-mode problem has addressed how to minimize picking and restocking costs in those warehouses with multiple forward modes, but the existing research assumes that all forward modes are restocked directly from the reserve mode. In our research, we will consider inventory systems where the forward modes can be restocked from storage modes other than the reserve mode. The cost of restocking a sku in a mode depends on the mode providing the restock.

5 Minimizing picking and restocking costs in the Avon inventory system

The first multi-tier inventory system that we examine is the Avon inventory system described below. We will show that this inventory system is equivalent to a two-tier system with two forward modes, meaning that we can then find a minimum cost slotting by using principles developed for multi-mode inventory systems in [1]. In the next section, we will extend these results to a general multi-tier inventory system.

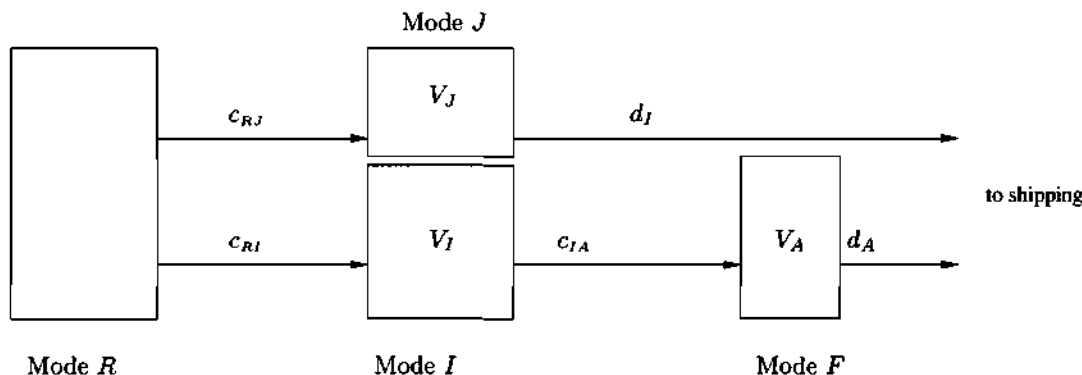


Figure 6: Configuration of Avon's warehouse in Suwanee, GA.

5.1 The Avon inventory system

The warehouse for Avon Products, Inc. outside Atlanta has a multi-tier inventory system. The company sells cosmetics and gift items; in 2002, they were the fifth biggest presence in the cosmetics industry by sales. They sell items through a network of 3.4 million sales representatives in 139 countries, each of whom collects the orders of their customers and places an order to a distribution center, or DC.

The warehouse outside fills orders for approximately 150,000 sales representatives in the southeastern United States. Sales representatives transmit their orders to the warehouse every two weeks, a period known as a *campaign*. Most representatives place only one order per campaign, but some with high sales volumes may place two or three. Between 150,000 and 170,000 orders will be processed in a typical two-week period. Each order requires 60 pieces, on average, meaning that over 10 million pieces are sold every two weeks. Because each representative usually requests only one or two of each sku they order, these 10 million pieces represent almost the same number of picks. Piece picking is therefore a very labor-intensive activity in the Avon warehouse, and it is a priority of the management to reduce costs for this as much as possible.

Orders are picked in a section of the warehouse called, appropriately, the *order fulfillment area*. The order fulfillment area has a footprint of approximately 120,000 sq. ft., and must have a picking location for each of 6,000-8,000 skus. Because there is limited storage space, most skus have additional supply in the bulk storage area of the warehouse, known as the *back warehouse*, and storage locations in the order fulfillment area are restocked

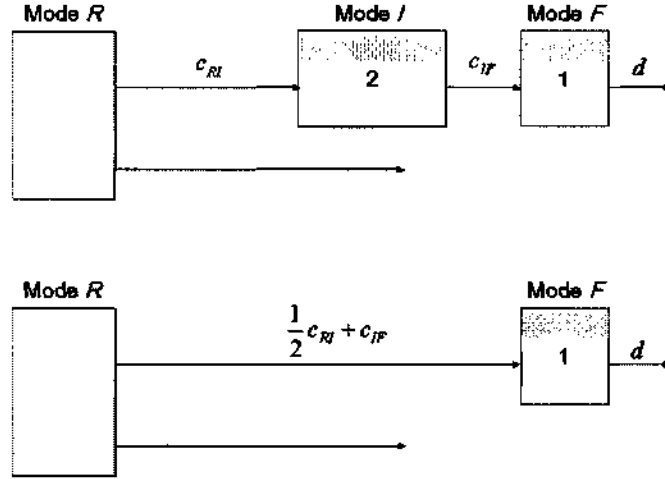


Figure 7: Example of equivalent inventory systems

from the back warehouse as needed. Restocks happen constantly as skus are being picked; one-third of the labor force in the order fulfillment area is dedicated to restocking skus.

A model of the Avon system is shown in Figure . Skus can be picked from the A-frame, denoted Mode A , or from a mode of shelving and flowrack which we call Mode J . All skus that are picked from Mode A have a supply in an intermediate Mode I , which also comprises several bays of flowrack. The two possible flowpaths in the system are thus R, I, A and R, J . It is cheaper to pick skus from the A-frame than from flowrack, so $d_A < d_J$.

5.2 Equivalence to a two-tier inventory system

To show the relationship between the Avon inventory system and a multi-mode system with two forward modes, we will make use of the following definition.

Definition 2 *Two inventory systems with the same number of flowpaths are equivalent if for each partition of the skus into flowgroups, the total cost of picking is the same in both systems and the cost of restocking is the same in both systems.*

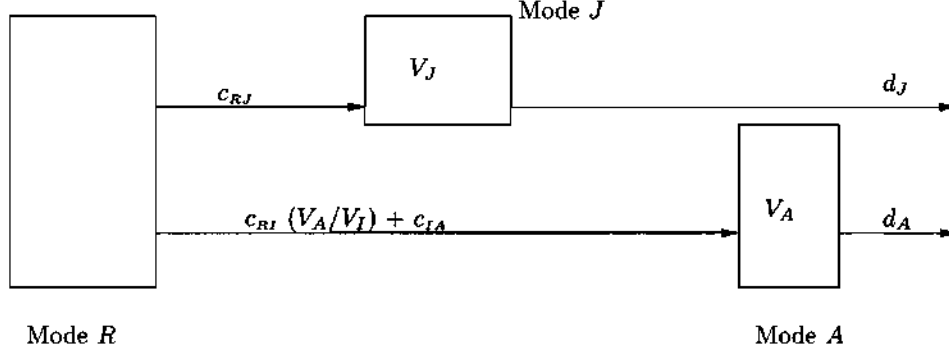


Figure 8: Inventory system with two forward modes that is equivalent to the Avon inventory system

By way of example, consider the top inventory system in Figure 7: one flowpath is R, I, F , and the other is R . Mode I has 2 cubic feet of storage space available, and Mode F has 1 cubic foot. Assume that skus have been partitioned into flowgroups. By results in [4] (Theorem 1), sku i flowgroup $S(R, I, F)$ will be allotted $2 \cdot \sqrt{f_i} \sum_{j \in S(R, I, F)} \sqrt{f_j}$ in Mode I , and $1 \cdot \sqrt{f_i} \sum_{j \in S(R, I, F)} \sqrt{f_j}$ in Mode F . Thus every time a sku in Mode F is restocked, half of the total capacity dedicated to that sku in Mode I is depleted. So for every restock of Mode F , we are charged for half a restock of Mode I . The total cost to the entire flowpath R, I, F each time the forward mode is restocked is then

$$\frac{1}{2}c_{RI} + c_{IF} \quad (2)$$

Now consider the bottom inventory system in Figure 7, with a reserve mode R' and a forward mode F' identical to Mode F . Each restock of Mode F' from Mode R' costs $c_{RI}/2 + c_{IF}$. Thus for any partition of a set of skus into flowgroups, the total cost of picking each flowgroup is the same in both the top and the bottom inventory systems, and the total cost of restocking each flowgroup also is the same. By our definition the forward-reserve inventory system on the bottom in figure 7 is equivalent to the one on the top.

Theorem 3 *The Avon inventory system has an equivalent inventory system of depth 2 with two forward modes.*

Proof Consider the Avon inventory system pictured in Figure 6. Assume that skus have been assigned to flowpaths and that space is allotted to the

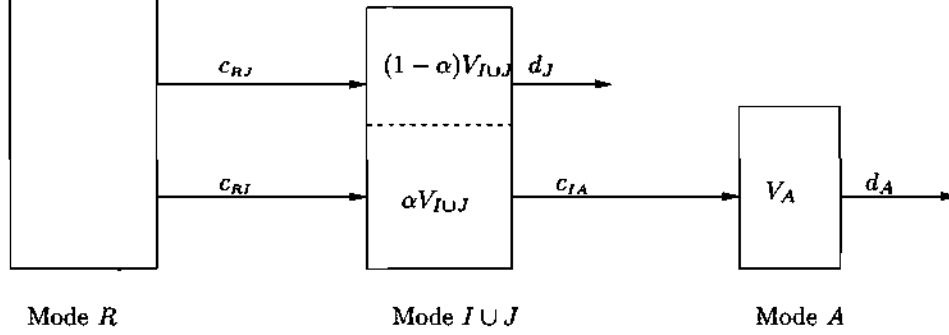


Figure 9: Avon system where all flowrack is considered one mode.

skus in each mode to minimize the cost of restocking the mode. By results in [4] (theorem 1), if sku i in flowpath R, I, A occupies volume v_{iA} in mode A , it will occupy $(V_I/V_A) \cdot v_{iA}$ in mode I . It follows that each restock of the sku in mode A contributes to V_A/V_I restocks of Mode I . The cost to the entire flowpath of one restock of mode A is thus $c_{IF} + c_{RI}(V_A/V_I)$.

If mode I were removed from the inventory system and sku i were restocked in mode A directly from the reserve mode at a cost per restock of $c_{IF} + c_{RI}(V_A/V_I)$, the total cost of restocking sku i in tier A would be unchanged. Thus for any partition of skus into flowgroups, the cost of restocking each flowgroup would be the same as in the Avon inventory system. Since the cost of picking is unchanged from the Avon inventory system, the two-tier inventory system formed by removing mode I and setting the cost of restocking mode A from mode I equal to $c_{IF} + c_{RI}(V_A/V_I)$ (pictured in Figure 8) is equivalent to the Avon inventory system. \square

Since we can effectively think of the Avon system as a two-tier inventory system with two forward modes, we can apply results from [1] (theorem 2) to find a near-minimum cost slotting of the Avon inventory system by considering only the $n+1$ slottings that correspond to partitioning the skus into two groups in descending order of viscosity.

5.3 Optimal allocation of storage resources

When visiting the Avon Products warehouse, we observed that the total supply of flowrack in the inventory system has been divided between Mode I and Mode J . We then wondered if each mode had received the proportion of flowrack that would minimize the cost of restocking the system.

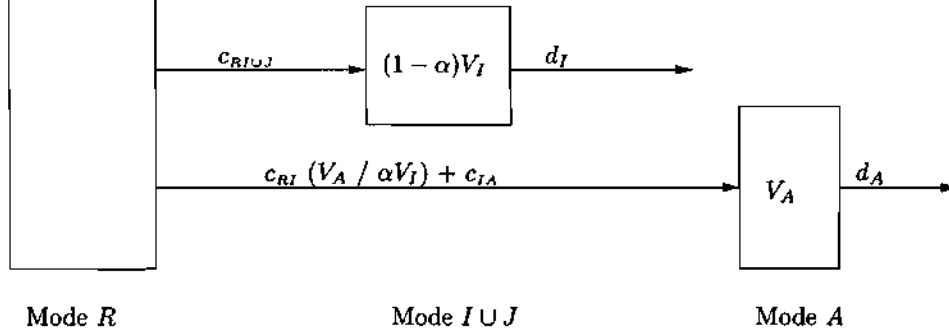


Figure 10: A two-mode system equivalent to the Avon system with one tier of flowrack.

To answer this question, we modeled the Avon inventory system as if the combined flowrack storage were one mode, Mode $I \cup J$, of volume $V_{I \cup J}$. We let the variable α represent the proportion of Mode $I \cup J$ dedicated to skus in $S(R, I, A)$, as shown in Figure 9. The optimal proportion of flowrack to be used for Mode I is the value of α in the minimum cost slotting, which we denote α^* .

The cost of restocking each flowpath in Mode $I \cup J$ remains as in the original Avon inventory system: the cost of restocking Mode A from Mode $I \cup J$ is c_{IA} , the cost per restock of a sku in flowpath R, I, A is c_{RI} , and the cost per restock of a sku in flowpath R, J is c_{RJ} . This assumption is reasonable if the optimal proportion of flowrack to dedicate to skus in each flowgroup is fairly close to the current proportion, since this means that restock costs in each area will remain approximately the same.

Finding a near-minimum cost slotting Let ϕ_q be the total rootflow assigned to flowpath q . For any assignment of skus to flowgroups in the combined-flowrack Avon inventory system, the cost of picking and restocking the inventory system is

$$\left(d_A \sum_{i \in S(R, I, A)} p_i + d_J \sum_{i \in S(R, J)} p_i \right) + c_{RI} \frac{\phi_{R, I, A}^2}{\alpha V_{I \cup J}} + c_{RJ} \frac{\phi_{R, J}^2}{(1 - \alpha) V_{I \cup J}} + c_{IA} \frac{\phi_{R, I, A}^2}{V_A} \quad (3)$$

The value of α that minimizes the cost of restocking Mode $I \cup J$, denoted α^* , is found by taking the derivative of Equation 3 with respect to α and

setting it equal to 0. Then

$$\alpha^* = \frac{\sqrt{c_{RI}} \cdot \phi_{R,I,A}}{\sqrt{c_{RI}} \cdot \phi_{R,I,A} + \sqrt{c_{RJ}} \cdot \phi_{R,J}}. \quad (4)$$

Substituting the expression for α^* in expression 4 into the cost function in 3, the cost of picking and restocking the system for a given assignment of skus to flowgroups can be written

$$d_A \sum_{i \in S(R,I,A)} p_i + d_J \sum_{i \in S(R,J)} p_i + \frac{(\sqrt{c_{RI}} \phi_{R,I,A} + \sqrt{c_{RJ}} \phi_{R,J})^2}{V_{I \cup J}} + c_{IA} \frac{\phi_{R,I,A}^2}{V_A} \quad (5)$$

Because the system is equivalent to the two-mode system shown in figure 10, we can find a near-minimum cost slotting by evaluating expression 5 for only the $n + 1$ assignments of skus to flowgroups that correspond to partitions of the skus ranked by viscosity. Once this slotting is identified, substitute the corresponding values of $\phi_{R,I,A}$ and $\phi_{R,J}$ into Expression 4 above to find the value of α^* , the optimal proportion of flowrack that should be dedicated to flowgroup $S(R, I, A)$.

6 Minimizing picking and restocking costs in general multi-tier inventory systems

In this section we show how to slot an arbitrary multi-tier inventory system to minimize total picking and restocking costs. We first show that any multi-tier inventory systems has the same fundamental structure as a two-tier inventory system with the same number of flowpaths. Consequently, we can find a near-optimal slotting for a multi-tier inventory system by using results developed for multi-mode systems in [3] and [1].

Theorem 4 *A multi-tier inventory system with P possible flowpaths is equivalent to a multi-mode system with P forward modes.*

Proof In a multi-tier system with P flowpaths, assume that skus have been partitioned into flowgroups. Consider a flowpath with Q tiers, numbered $1, \dots, Q$, where skus assigned to the flowpath are picked from tier Q . (See figure 11.) Let the space allotted to the flowgroup in tier q be denoted V_q , and let each restock of tier q from its predecessor tier cost c_q . By results in [4] (theorem 1), if sku i in flowpath occupies volume $v_{i,Q}$ in tier Q , it

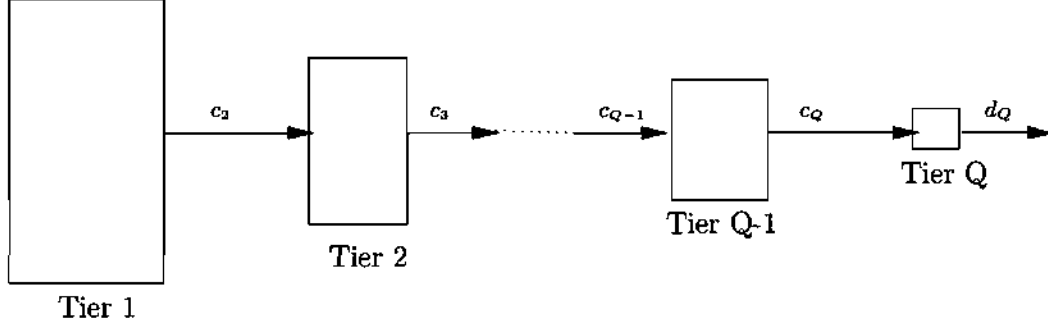


Figure 11: Flowpath with Q tiers.

will occupy $(V_q/V_Q) \cdot v_{iQ}$ in tier q . Thus each restock of a sku in tier Q contributes to (V_Q/V_q) restocks of the sku in tier q . The total cost to the flowpath for each restock of the flowgroup in tier Q is then

$$\sum_{q=1}^Q c_q \frac{V_Q}{V_q} \quad (6)$$

The total cost per restock of the flowgroup in tier Q is the same as if we restock a sku in the flowgroup in tier Q directly from the reserve mode with a cost per restock as given in Expression 6. Since this holds for all P flowpaths, the multi-tier system is equivalent to a multi-mode system with P forward modes. \square

Figure 12 shows a multi-tier inventory system with three flowpaths; if we assume that flowpath C occupies $\alpha_2 V_2$ cubic feet in Mode 2 and $\beta_3 V_3$ cubic feet in Mode 3, then Figure 13 is an equivalent system.

One important result of this equivalence is that we can find a near-optimal slotting of a multi-tier inventory system by considering only those slottings that correspond to a partition of the skus by viscosity. The fact that we must consider only such slottings reduces the number of slottings we need to consider when solving the problem by enumeration, but with several modes and thousands of skus, finding a near-optimal solution by enumeration can be computationally difficult. Bartholdi and Hackman present a method for finding the near-optimal slotting of a multi-mode inventory system in $O(\log n^M)$ time in [1]. We can use this result to efficiently find the near-optimal slotting of a multi-tier inventory system if no flowgroups share a mode in the system: we need only derive the equivalent multi-mode

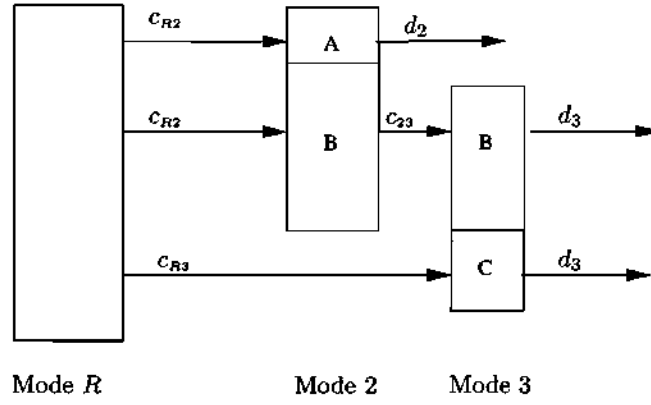


Figure 12: Multi-tier inventory system.

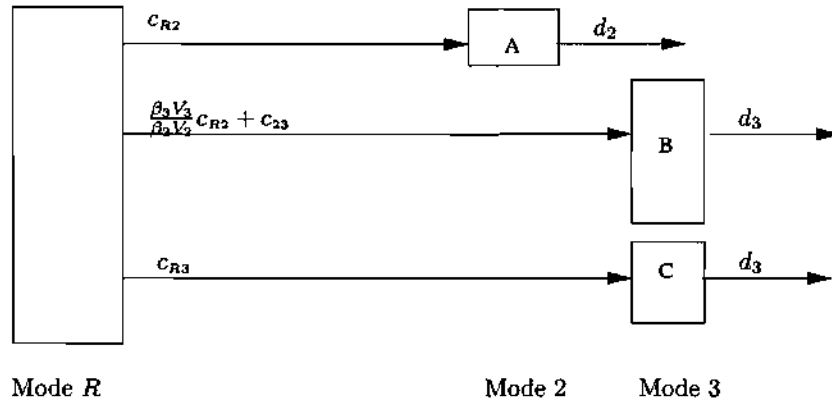


Figure 13: Multi-mode inventory system equivalent to system in Figure 12.

inventory system and apply the technique.

If the multi-tier inventory system has a mode that is shared by flowgroups, however, then the cost to restock a flowgroup in the equivalent two-tier inventory system depends on the amount of space that the flowgroup is allotted in each shared mode. This amount of space depends on the total amount of rootflow assigned to each flowgroup, meaning that the restock costs for the equivalent two-tier system will be different for each slotting that we consider. Because of this, we cannot directly apply the faster search method derived in [1]. A goal for future research is to identify properties of multi-tier inventory systems with shared modes that allow us to either adapt the search method in [1] or to develop similar search methods to reduce the number of potential slottings we must consider to find a near-optimal solution.

References

- [1] John J. Bartholdi, III and Steven T. Hackman. Simple ranking properties for the multi-mode forward-reserve problem. Working paper.
- [2] Edward H. Frazelle. *World-Class Warehousing*. Logistics Resources International, Inc., Atlanta, 1996.
- [3] Steven T. Hackman and Loren K. Platzman. Near-optimal solution of generalized resource allocation problems with large capacities. *Operations Research*, 38(5):902-910, September-October 1990.
- [4] Steven T. Hackman and Meir J. Rosenblatt. Allocating items to an automated storage and retrieval system. *IIE Transactions*, 22(1):7-14, March 1988.